Zvonimir Šikić

THE TWELVE MAGNIFICENTS

Abstract. We present the history of the theory and practice of musical scales, which culminated in today's universal 12-tone scale. Then we prove that this 12-part division of the octave is optimal. The proof is an application of the theory of continued fractions. Drobisch in *Über musikalische Tonbestimmung und Temperatur* (1852) used continued fractions for subdividing the octave into intervals. It is not clear whether he was the first to use this approach; e.g. Euler was dealing with mathematics of temperament and also wrote a tract on continued fractions, *De Fractionibus Continuis Dissertatio* (1737).

Mathematics Subject Classification (2020): 00A65

Keywords: interval, scale, tuning, temperament, continued fraction

Centar za logiku i teoriju odlučivanja Sveučilišta u Rijeci
zvonimir@sikic.com
ORCID: 0000-0002-9698-6032

DOI: https://doi.org/10.18485/mi_sanu_zr.2024.29.21.ch2

Contents

1. Intervals and scales	24
2. Equal vs. just	28
3. Pythagorean tuning	31
4. Why exactly twelve semitones?	34
References	39

1. Intervals and scales

The concept of interval and the concept of scale derived from it are extremely important in the theory of music. They are also crucial for understanding historical development of music. Let us imagine a segment of piano keys that is periodically repeated on the piano and that contains typical groups of 2 and 3 black keys within 8 white keys.

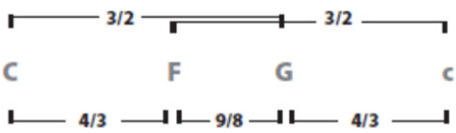
A series of 8 "white" tones C, D, E, F, G, A, H, c (the so-called diatonic scale) is the well-known major scale, in this case C major. The octave interval, from C to c, contains 8 "white" tones. The interval of a fifth, from C to G, contains 5 "white" tones. The interval of a fourth, from C to F, contains 4 "white" tones. With the agreement that the frequency of tone C is unity, we have the following frequencies:

$$\mathrm{C} = 1, \quad \mathrm{F} = 4/3, \mathrm{G} = 3/2 \quad \text{and} \quad \mathrm{c} = 2.$$

As the size of the intervals between two tones is equal to the ratio of their frequencies, it is easy to calculate the sizes of the following intervals:

$$\frac{\mathrm{c}}{\mathrm{F}} = \frac{2}{\frac{4}{3}} = \frac{3}{2} \qquad \frac{\mathrm{c}}{\mathrm{G}} = \frac{2}{\frac{3}{2}} = \frac{4}{3} \qquad \frac{\mathrm{G}}{\mathrm{F}} = \frac{\frac{3}{2}}{\frac{4}{3}} = \frac{9}{8}$$

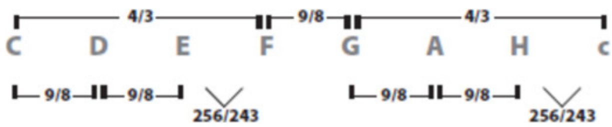
Hence, c is the fifth on F and c is a fourth on G. The interval from F to G, of size 9/8 (or 204 cents, see below), is called a whole tone or major second. So, the fourth (with the frequency 4/3) and the fifth (with the frequency 3/2) form the following division of the octave:



If we divide the lower fourth from C to F, and the upper fourth from G to c, by whole tones of length 9/8, we will get two more tones in each of them:

D/C = 9/8, i.e. due to C = 1, D = 9/8.
E/D = 9/8, i.e. due to D = 9/8, E = 81/64.
A/G = 9/8, i.e. due to G = 3/2, A = 27/16.
H/A = 9/8, i.e. due to A = 27/16, H = 243/128.

Thus we arrive at the following division of the octave:



The interval from E to F, like the interval from H to c, has a length of 256/243, because

$$\frac{\frac{4}{3}}{\left(\frac{9}{8}\right)^2} = \frac{256}{243}.$$

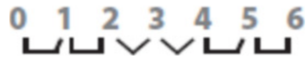
This interval (which has 90 cents, see below) is called a diatonic semitone or limma. The whole tones and semitones obtained in this way have the bad property that two semitones are smaller than one whole tone:

$$\left(\frac{256}{243}\right)^2 = 1.110 < 1.125 = \frac{9}{8}.$$

The octave can be divided in other ways. The most famous are the modes or modal scales that can be played on the white keys of the piano, starting from different basic tones. The Greeks used several modes: the Ionian mode based on C (which is our C major), the Doric mode based on D, the Phrygian mode based on E, the Lydian mode based on F, the Mixolydian mode based on G, the Aeolian mode based on A and the Locrian mode based on H. Of course, the octave can be divided into scales that are not "white". Such are, for example, ascending and descending melodic minor.

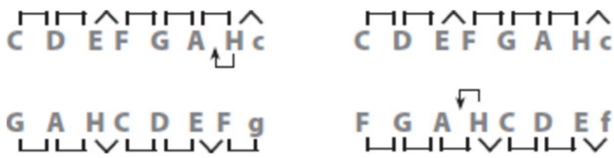
Pentatonic scales, i.e. scales that divide the octave into 5 intervals (as opposed to the so far considered heptatonic scales that divide it into 7 intervals), usually insert one tone between the tonic and the fourth, and another between the fifth and the higher tonic. Although the term pentatonic scale can be used for any division of the octave into 5 intervals, in Europe the term has become established for heptatonic "white" scales from which the "semitones" F and H are omitted.

Of the hexatonic scales, which divide the octave into 6 intervals, the most famous are the whole-tone scale, which divides the octave into 6 whole tones, and the blues scale, which divides it as follows (with the obvious markings for a whole tone, semitone, and one and a half tones):



The history and distribution of these scales are complex. The diatonic division of the octave (into "white" tones), which led to the heptatonic modes, goes back to the time of Pythagoras. The use of the lyre and the organ (invented around 300 BC by Ctesibius of Alexandria) both of which had D as their main tone, led to the predominance of the Doric mode in musical antiquity. In the early Middle Ages, within the scholae cantorum, the Christian church continued the tradition of monophony (performance of only one melodic line) in all ancient modes. Gregory's reform of this school led to the Gregorian chorale, which is still recognizable today for its use of ancient modes.

At the turn of the 10^{th} to the 11^{th} century, experiments with polyphony began - the simultaneous performance of several melodic lines. The voices were arranged in parallel motion, so that one plays his melody while the other follows, simultaneously playing that same melody a fourth or fifth lower. So, the tones of the first voice: C D E F G A H c correspond to the tones of the second voice: G A H C D E F g (which are a fourth lower) or F G A H C D E f (which are a fifth lower). The voices were always separated by a fourth (2.5 whole tones) or a fifth (3.5 whole tones), except when they simultaneously sing the tritone H/F (3 whole tones). The tritone is extremely dissonant and difficult to sing. That is why when the tones H and F meet, there would always be an uncertainty. It could be eliminated by lowering the tone H by one semitone.

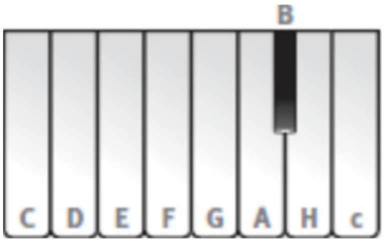


Thus, the old modes were expanded with the note B, and the first black key was inserted into the keyboards. It was the only flattened note and the only black key for a while.

We can notice that the shift from H to B in the first voice cancels the leading tone (the final semitone interval that naturally closes the musical phrase). That is why a shift from F to $F^{\#}$ in the second voice seems to be a better solution. This shift removes the dissonant tritone, but does not remove the leading tone. Of course, the new notes B and $F^{\#}$ naturally lead to new fifths:

$$E^b \leftarrow B \qquad F^\sharp \rightarrow C^\sharp \rightarrow G^\sharp.$$

Thus, the octave was finally divided into 12 semitones, and the keyboards got 4 more black keys within the octave:



This process did not directly lead to twelve-note composition in the obtained dodecatonic (chromatic) scale. That development occurred only in the 20^{th} century. In this sense, the division into semitones did not become a scale for a long time. It only enabled polyphonic composition in the old heptatonic modes, because in it any old scale could be attached to any tone as a tonic (in other words, it allowed for transpositions). This also opens the door to harmonic composition, which actually has no clear boundaries with polyphonic, since the simultaneous (vertical) aspect of polyphony is actually harmony.

Expanding music into two dimensions created a vast range of possibilities. There were so many great composers in Europe in the 16^{th} century that one lifetime is not enough to get to know even a part of their oeuvre. Students of polyphony, i.e. counterpoint, still learn from that inexhaustible source. Nevertheless, after countless experiments with all traditional modes, at the beginning of the 17^{th} century, Ionic mode stood out. That is today's major. The other modes slowly disappeared, except for the modified variant of Aeolian mode, which is today's minor. Both surviving modes are characterized by a major seventh, the so-called leading tone, which is only a semitone away from the tonic, thus strongly emphasizing it as the resolution point for its dissonant tension. The leading tone infatuation with the tonic is the main characteristic of the surviving modes. From Baroque to Romanticism, they undisputedly ruled Western tonal, which means tonic, music.

Composers from the beginning of the 20^{th} century faced new doubts. Richard Wagner's *Unendliche Melodie*, which is actually a melody without a tonic end, and its "wandering" chords, brought tonal music to the very limits of tonality (i.e. tonicity). Wagner's immediate followers were no longer satisfied with the old system of tonality and its relationships. They were looking for new scales without a conspicuous tonic. Debussy composed in a whole-tone scale without a conspicuous tonic, which gives his music a calm, almost otherworldly quality. The same scale was used almost regularly by jazz giant Thelonius Monk, and it is still extremely popular in many jazz circles. Schönberg and similar 20^{th} century composers used the chromatic scale, with effects often foreign to the unaccustomed ear.

However, there is also a noticeable return to traditional heptatonic modes, which are old enough and forgotten enough to appear again as a fresh and new sound. Like the sound of old, almost primitive times, or like the timeless sound of Gregorian chants. At the beginning of his 6^{th} symphony, Sibelius takes us to the old Finnish forests using the Doric mode. Debussy, Hindemith and Stravinsky often use the Doric absence of the leading tone to achieve the effect of a solemn, almost formal A-MEN (which is a whole-tone ending).

The same effect is achieved with the Phrygian mode. Its additional characteristic is the sad and oriental sound of its semitone beginning. That is why Rimsky-Korsakov used it in his *Scheherazade* and Brahms in the slow movement of his 4^{th} symphony. Much of the "sad" Spanish, Gypsy and Jewish music was composed in the Phrygian mode, e.g. Liszt's 2^{nd} Hungarian (actually Gypsy) Rhapsody.

The augmented fourth of the Lydian mode seems like a wrong note in the common major. It almost sounds like a joke. This is precisely why the witty Prokofiev often used the Lydian mode. Otherwise, this mode is characteristic of Polish music; both of folk music and of Chopin's polonaises and mazurkas. The "Polonaise" from Act 3 of Mussorgsky's *Boris Godunov* was written in the Lydian mode.

The Mixolydian mode is "major with diminished seventh", i.e. "major without a leading tone". This made it unusual and innovative enough to allow it to enter many rock-and-roll classics (Kinks' *You really got me*, Beatles' *Norwegian wood*, etc.).

The pentatonic scale C D E G A c (also called "black" scale, because starting from $F^{\#}$, can be played on the black keys as $F^{\#}$ $G^{\#}$ B $C^{\#}$ $D^{\#}$) is used in folk music worldwide. If you go through the pentatonic sequence C D E G A c several times (in arpeggio), at least one of the popular music hits will appear in your mind. The entertainment industry has profited a lot from these five notes. It is the most common popular music scale. It solves the H/F tritone problem by removing the notes H and F from the heptatonic scale: C D E F G A H c. We could therefore call it a minimalist solution for polyphonic and harmonic composing.

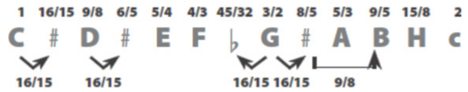
2. Equal vs. just

Let us go back to the basic musical material consisting of 12 semitones of the chromatic scale (from which all the scales discussed in the previous chapter choose their material):

So far, we have not thoroughly studied the problem of the size of the intervals in that scale. The Pythagorean ideal would be, basic tone C = 1/1, octave c = 2/1, fifth G = 3/2, fourth F = 4/3 and third E = 5/4. It follows:

$$\frac{\mathrm{G}}{\mathrm{F}} = \frac{\frac{3}{2}}{\frac{4}{3}} = \frac{9}{8} \qquad \frac{\mathrm{F}}{\mathrm{E}} = \frac{\frac{4}{3}}{\frac{5}{4}} = \frac{16}{15} \qquad \frac{\mathrm{D}}{\mathrm{C}} = \frac{9}{8} \qquad \frac{\mathrm{E}}{\mathrm{D}} = \frac{\frac{5}{4}}{\frac{9}{8}} = \frac{10}{9}.$$

That scale is recommended by Ptolemy, the greatest ancient astronomer and music theorist. Its semitones are 16/15 in length, while the whole tones are of different lengths, 9/8 and 10/9 (namely, E/D = (5/4)/(9/8) = A/G = (5/3)/(3/2) = 10/9). By breaking whole tones into semitones, we finally arrive at all the just intervals in the twelve-tone scale:



Namely,

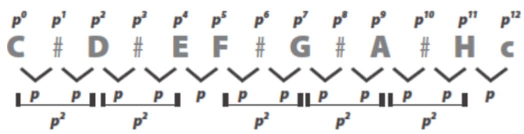
$$\mathrm{C}^\sharp = \mathrm{C}\cdot\frac{16}{15} = 1\cdot\frac{16}{15} = \frac{16}{15} \qquad \mathrm{D}^\sharp = \mathrm{D}\cdot\frac{16}{15} = \frac{9}{8}\cdot\frac{16}{15} = \frac{6}{5}\mathrm{G} = \mathrm{G}\cdot\frac{16}{15} = \frac{3}{2}\cdot\frac{16}{15} = \frac{8}{5}$$

$$\mathrm{G}^\mathrm{b} = \mathrm{G} : \frac{16}{15} = \frac{3}{2} : \frac{16}{15} = \frac{45}{32} \qquad \mathrm{B} = \mathrm{G}^\sharp\cdot\frac{9}{8} = \frac{8}{5}\cdot\frac{9}{8} = \frac{9}{8}.$$

The magnitudes of the semitones of the chromatic scale thus obtained, in addition to the value 16/15 = 1.07, have the values 25/24 = 1.04, 135/128 = 1.05 and 27/25 = 1.08 (unmarked semitones have the value 16/15):

These Pythagorean just intervals are part of the so-called natural intonation, so musical instruments that are tuned in this way are said to be tuned in natural intonation. The basic problem of natural intonation is a large number of different semitones, even four, which often makes faithful transpositions impossible. This problem is completely solved by the equally tempered scale, i.e. equal temperament recommended by the greatest ancient music theorist Aristoxenus.

The equally tempered scale is the division of the octave into 12 equal semitones. If we mark the semitone interval with p, it is the following division:



Of course, $\mathrm{p} = 2^{1/12} \approx 1.06$. The distance of the tone from the tonic is more often denoted by the exponent (then by the whole power). It is a transition from the multiplicative scale $(\mathrm{p}^0, \mathrm{p}^1, \mathrm{p}^2, \dots)$ to the additive logarithmic scale $(0, 1, 2, \dots)$. This allows us to add the intervals, instead of multiplying them (e.g. the fact that the third and the fifth make a seventh, in the logarithmic scale we find by addition, $4 + 7 = 11$, and in the basic one by multiplication, $\mathrm{p}^4 \cdot \mathrm{p}^7 = \mathrm{p}^{11}$).

For a more precise determination of the intervals, a smaller unit cent is introduced. It is a hundredth part of a semitone, i.e. s = p/100, so s = $2^{1/1200} \approx 1.00058$. This means that in the logarithmic scale expressed in cents (i.e. in base s):

minor second	=	100 cent	fifth	=	700 cent
major second	=	200 cent	minor sixth	=	800 cent
minor third	=	300 cent	major sixth	=	900 cent
major third	=	400 cent	minor seventh	=	1000 cent
forth	=	500 cent	major seventh	=	1100 cent
tritone	=	600 cent	octave	=	1200 cent

If we want to express an interval of length m/n in cents, we must actually calculate the logarithm of m/n in base s:

$$\log_s \frac{m}{n} = \frac{\log \frac{m}{n}}{\log_s} = 1200 \frac{\log \frac{m}{n}}{\log 2}$$

With the help of this formula, we can express the just intervals in cents and compare them with equally tempered ones:

16/15	=	minor second	=	111.73 cent	3/2	=	fifth	=	702.00 cent
9/8	=	major second	=	203.91 cent	8/5	=	minor sixth	=	813.69 cent
6/5	=	minor third	=	315.64 cent	5/3	=	major sixth	=	884.36 cent
5/4	=	major third	=	386.31 cent	9/5	=	minor seventh	=	1017.60 cent
4/3	=	forth	=	498.00 cent	15/8	=	major seventh	=	1088.27 cent
45/32	=	tritone	=	590.22 cent	2/1	=	octave	=	1200.00 cent

Fourths, fifths, and seconds are almost the same in just and equally tempered scale (a deviation of up to 10 cents is not large), but thirds, sixths, and sevenths show larger deviations. It is obvious that these are essentially different tunings of the "same" tones. The main advantage of natural intonation, its accuracy, is lost when transposing. Compare the equal temperament shift from C to the thirds of Eb and E and to the fifth of G, with the same shifts in natural intonation:

C	♭	D	♭	E	F	♭	G	♭	A	B	H	c	♭	d	♭	e	f	♭	g
0	100	200	300	400	500	600	700	800	900	1000	1100	1200	1300	1400	1500	1600	1700	1800	1900
			0	100	200	300	400	500	600	700	800	900	1000	1100	1200				
				0	100	200	300	400	500	600	700	800	900	1000	1100	1200			
							0	100	200	300	400	500	600	700	800	900	1000	1100	1200

C	#	D	#	E	F	♭	G	#	A	B	H	c	#	d	#	e	f	♭	g
0	112	204	316	386	498	590	702	814	884	1018	1088	1200	1312	1404	1516	1586	1698	1790	1902
			0	70	182	274	386	498	568	702	772	884	996	1088	1200				
				0	112	204	316	428	498	632	702	814	926	1018	1130	1200			
							0	112	182	316	386	498	610	702	814	884	996	1088	1200

All shifts in the equally temperd scale give exactly the same series of semitones $(100, 200, \ldots, 1200)$, while shifts in the natural scale have 3 errors in the shift of the fifth, 4 errors in the shift of the major third, and even 6 errors in the shift of the minor third (errors are rounded in the table). These mistakes are not small. Of the 13 rounded, 7 deviate from the correct values by 42 cents, and 6 by 22 cents ("tolerable" error is up to 10 cents).

Errors of the equal temperament, which are not altered by transpositions, range from 0 cents to 18 cents, with an average of exactly 10 cents.

However, music history did not follow the path of music theory. It is theoretically justified to think about Ptolemy's intervals, but it is difficult to realize them on actual musical instruments. Who can correctly sing (or tune) two different major seconds 9/8 and 10/9, minor second 16/15 or $2^{1/12}$, $2^{2/12}$, $2^{3/12}$ etc.?

On the other hand, it is easiest to sing and recognize perfect octaves and fifths, which is why throughout antiquity and the Middle Ages the standard tuning was the Pythagorean tuning with fifths and octaves.

3. Pythagorean tuning

In the 13th century, the French Academy of Notre Dame declared that the correct scale can only be reached by a series of perfect Pythagorean fifths, in which the constant ratio to the previous tone is the "divine" ratio of 3:2 (3 for the Holy Trinity, 2 for various dualisms; heaven and earth, spirit and body, good and evil, etc.).

We start from the tone $C = 1$ as a tone of unit frequency. A fifth above is $G = 3/2$. A fifth above $G = 3/2$ is $d = (3/2)^2$. That tone goes outside the octave that extends from $C = 1$ to $c = 2$, so we lower d by an octave to get $D = (1/2)(3/2)^2 = 9/8$. A fifth higher is the tone $A = (1/2)(3/2)^3 = 27/16$. The fifth above A is $e = (1/2)(3/2)^4 = 81/32$, which we lower again by an octave to get $E = (1/2)2(3/2)^4 = 81/64$. A fifth above E is $H = (1/2)^2(3/2)^5 = 243/128$. That is how we got all the "white" tones except F. Let us down a fifth below the initial $C = 1$ and we'll get $F' = 1/(3/2) = 2/3$, so we can easily find an octave higher $F = 2(2/3) = 4/3$. Thus, we tuned all the "white" tones of the diatonic scale using only fifths and octaves:

$$F \leftarrow C \rightarrow G \rightarrow D \rightarrow A \rightarrow E \rightarrow H$$

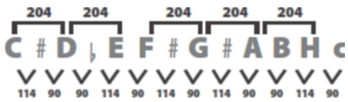
Let us go down from F by two fifths and we get (with corresponding shifts to basic octave) $B = 2^2(2/3)^2 = 16/9$ and $E^b = 2^2(2/3)^3 = 32/27$. Let us go up three fifths from H and (with corresponding shifts to basic octave) we have $F^\sharp = (1/2)^3(3/2)^6 = 729/512$, $C^\sharp = (1/2)^4(3/2)^7 = 2187/2048$ and $G^\sharp = (1/2)^4(3/2)^8 = 6561/4096$. This is how we tuned all the tones of the chromatic scale:

$$E^b \leftarrow B \leftarrow F \leftarrow C \rightarrow G \rightarrow D \rightarrow A \rightarrow E \rightarrow H \rightarrow F^\sharp \rightarrow C^\sharp \rightarrow G^\sharp$$

This Pythagorean tuning is also called all-fifths tuning or the circle of fifths. In this case it is a Pythagorean E^b - $G^\#$ tuning, because from C we move down to E^b and up to $G^\#$. The obtained values can be expressed as ratios or more clearly in logarithmically calculated cents (values in cents are rounded up to integer values):

C	#	D	♭	E	F	#	G	#	A	B	H	c
1/1	2187/2048	9/8	32/27	81/64	4/3	729/512	3/2	6561/4096	27/16	16/9	243/128	2/1
0	114	204	294	408	498	612	702	792	906	996	1110	1200

Notice that the semitones of the scale thus obtained are not equal, they amount to 114 cents or 90 cents. The smaller one is the Pythagorean diatonic semitone also known as limma (the ratio 256/243 has 90 cents). The larger one is known as the Pythagorean apotome.



We see that a whole tone can be 204 cents or 180 cents. But since the 114-cent apotome always occurs between the tone X and its sharpened $X^{\#}$ or flattened X^{b}, it follows that each whole tone of the diatonic "white" scale has 204 cents.

This tuning has one major drawback. All the fifths we tuned consist of 4 diatonic semitones and 3 apotomes, $4\cdot90+3\cdot114 = 702$, and they are perfect. Unfortunately, the fifth $G^{\#}$ - e^{b}, which is the end result of our tuning, and which is not directly tuned, consists of 5 diatonic semitones and 2 apotomes, $5 \cdot 90 + 2 \cdot 114 = 678$. It does not sound good and is called a wolf interval, because it sounds like a wolf howl. The same goes for the E^{b} - $G^{\#}$ fourth, which is $3 \cdot 114 + 2 \cdot 90 = 522$, instead of the perfect $3 \cdot 90 + 2 \cdot 114 = 498$. Any Pythagorean tuning will have wolf intervals like these. They are determined by the final notes in the sequence of fifths, because 12 fifths that cover a range of 7 octaves (and at the same time "hit" all the tones of the chromatic scale) do not do so exactly. Namely, $2^7 < (3/2)^{12}$ which is why E^{b} and $G^{\#}$ do not form perfect, but wolf intervals. Expressed in cents (an octave has 1200, and a perfect fifth has 702):

$$7 \cdot 1200 < 12 \cdot 702 \qquad \text{i.e.} \qquad 7 \cdot 1200 < 12 \cdot 700 + 12 \cdot 2$$

This means that wolf intervals differ from perfect intervals by $12 \cdot 2 = 24$ cents. The wolf fourth is 24 cents too big, and the wolf fifth is 24 cents too small. The interval of 24 cents, which makes these differences, is called the Pythagorean comma. (Note that it is equal to the difference of the Pythagorean semitones of 114 cents and 90 cents, which is no coincidence. This and many other regularities are explained in [2].)

In addition to one imperfect fifth and fourth, Pythagorean tuning also generates thirds, all of which are imperfect. The major third of this tuning has a length expressed by the ratio 81/64, which differs from the perfect $5/4 = 80/64$. The interval, for which these major thirds differ, has a length expressed by the ratio $(81/64)/(80/64) = 81/80$, which is 22 cents. The same interval separates the perfect minor third 6/5 from the tuned 32/27, because $(6/5)/(32/27) = 81/80$. So, a major third tuned by all-fifths tuning is 22 cents higher than perfect, while a minor third tuned by all-fifths tuning is 22 cents lower than perfect. Their sum is a perfect fifth (because the excess and deficiency of 22 cents cancel out), except on

E^b and $G^{\#}$, where we have almost perfect minor thirds, but (therefore) bad major thirds. The interval of 22 cents, by which perfect thirds differ from tuned, is called a syntonic comma.

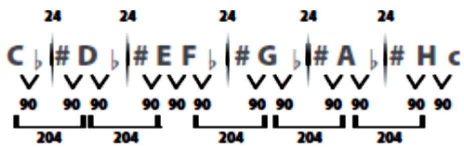
The syntonic comma is characteristic of all thirds in the Pythagorean tuning system, while the Pythagorean comma is characteristic of only one wolf fifth and only one wolf fourth. That is why Pythagorean tuning is particularly suitable for performing music in which fourths and fifths are the dominant consonances (with the necessary avoidance of the wolfish E^b and $G^{\#}$) and in which thirds form a dissonant element. Such is the case with Gothic polyphony, with its actively dissonant thirds, perfect fifths and fourths, and a small diatonic semitone (from only 90 cents) for effective cadences. On the other hand, the harmony of triads based on thirds and fifths, which appeared in the 15^{th} century and rules Western music to this day, cannot tolerate the syntonic comma and its bad thirds. That is why musicians from the Renaissance to Romanticism will use new tunings, which eliminate the syntonic comma. Let us also mention that Pythagorean G^b - H tuning, which uses the following sequence of fifths:

$$G^b \leftarrow D^b \leftarrow A^b \leftarrow E^b \leftarrow H^b \leftarrow F \leftarrow C \rightarrow G \rightarrow D \rightarrow A \rightarrow E \rightarrow H$$

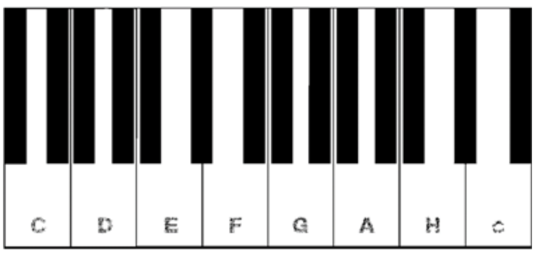
and F - $A^{\#}$ tuning, which uses the sequence:

$$F \leftarrow C \rightarrow G \rightarrow D \rightarrow A \rightarrow E \rightarrow H \rightarrow C^\sharp \rightarrow G^\sharp \rightarrow D^\sharp \rightarrow A^\sharp,$$

generate a total of 17 tones with the help of which we can also perform perfect thirds:



This 17-tone scale has been known since the 15^{th} century, and there were also keyboards that realized it with double black keys. They enabled performance of flattened tones b and sharpened #, which are identical in the equal temperament, but not in the Pythagorean one.



4. Why exactly twelve semitones?

Why does an octave contain 12 semitones? We have seen that the pentatonic, hexatonic and heptatonic scales choose their tones from the basic dodecatonic division, but we also encountered the #b-distribution which has 17 tones in the octave. Microtone composers of the 20^{th} century use a scale with 41 microtones and many other numbers, but the supremacy of the magnificent 12 is still unquestionable.

We have seen how the Pythagorean exact intervals 1:2:3:4:5 historically led to the dominance of the twelve-tone octave division. Now let us see what mathematics leads us to, if we give priority to Pythagoras' small ratios.

Let us choose the first tone for our music and consider its frequency as unity. Despite Jobim's *Samba por una nota sô*, one note is not enough. In accordance with Pythagoras, let us add to our unit tone all those tones that are a certain number of octaves below and above it. In the range of frequencies that we hear (from 20 Hz to 20,000 Hz), it is a maximum of 10 octaves:

$$2^{-5} \leftarrow 2^{-4} \leftarrow 2^{-3} \leftarrow 2^{-2} \leftarrow 2^{-1} \leftarrow 1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5.$$

There are rare voices that can move beyond a range of two octaves, which means that they currently have a maximum of three tones at their disposal, which in addition sound the same. Therefore, in accordance with Pythagoras and the proclamation of the academy of *Notre Dame*, we will add fifths and their corresponding octaves to our sheet music:

$$\begin{gathered}
\frac{1}{2^2}\left(\frac{3}{2}\right)^{-3} \leftarrow \frac{1}{2^2}\left(\frac{3}{2}\right)^{-2} \leftarrow \frac{1}{2^2}\left(\frac{3}{2}\right)^{-1} \leftarrow \frac{1}{2^2} \rightarrow \frac{1}{2^2}\left(\frac{3}{2}\right)^{1} \rightarrow \frac{1}{2^2}\left(\frac{3}{2}\right)^{2} \rightarrow \frac{1}{2^2}\left(\frac{3}{2}\right)^{3} \\
\frac{1}{2}\left(\frac{3}{2}\right)^{-3} \leftarrow \frac{1}{2}\left(\frac{3}{2}\right)^{-2} \leftarrow \frac{1}{2}\left(\frac{3}{2}\right)^{-1} \leftarrow \frac{1}{2} \rightarrow \frac{1}{2}\left(\frac{3}{2}\right)^{1} \rightarrow \frac{1}{2}\left(\frac{3}{2}\right)^{2} \rightarrow \frac{1}{2}\left(\frac{3}{2}\right)^{3} \\
\left(\frac{3}{2}\right)^{-3} \leftarrow \left(\frac{3}{2}\right)^{-2} \leftarrow \left(\frac{3}{2}\right)^{-1} \leftarrow 1\left(\frac{3}{2}\right)^{1} \rightarrow \left(\frac{3}{2}\right)^{2} \rightarrow \left(\frac{3}{2}\right)^{3} \\
2\left(\frac{3}{2}\right)^{-3} \leftarrow 2\left(\frac{3}{2}\right)^{-2} \leftarrow 2\left(\frac{3}{2}\right)^{-1} \leftarrow 2 \rightarrow 2\left(\frac{3}{2}\right)^{1} \rightarrow 2\left(\frac{3}{2}\right)^{2} \rightarrow 2\left(\frac{3}{2}\right)^{3} \\
2^2\left(\frac{3}{2}\right)^{-3} \leftarrow 2^2\left(\frac{3}{2}\right)^{-2} \leftarrow 2^2\left(\frac{3}{2}\right)^{-1} \leftarrow 2^2 \rightarrow 2^2\left(\frac{3}{2}\right)^{1} \rightarrow 2^2\left(\frac{3}{2}\right)^{2} \rightarrow 2^2\left(\frac{3}{2}\right)^{3}
\end{gathered}$$

In this way, we arrive at an infinite number of tones, within each octave, which are given by the ratios of the form:

$$\frac{1}{2^m}\left(\frac{3}{2}\right)^n \quad m, n = 0, \pm 1, \pm 2, \ldots$$

For each exponent n, we have one new tone, which is shifted m octaves lower or higher using exponents m. None of those tones match the base tone, because it is

$$\frac{1}{2^m}\left(\frac{3}{2}\right)^n \neq 0, \quad 3^n \neq 2^{m+n}.$$

Hence, the procedure of the Notre Dame academy constantly generates new tones. Music with an infinite number of notes is just as impossible as music with only one, so we have to stop the generation process somewhere. It is natural to stop at the n-th tone if, returned to the basic octave, it at least approximately coincides with the basic tone (since exact matching is impossible). For example, for the 7^{th} and 12^{th} tones we have the following approximations:

$$\frac{1}{2^4}\left(\frac{3}{2}\right)^7 \approx 1 \quad \text{i.e.} \quad \frac{3}{2} \approx 2^{\frac{4}{7}} = 1.486 \text{ (686c)}$$
$$\frac{1}{2^7}\left(\frac{3}{2}\right)^{12} \approx 1 \quad \text{i.e.} \quad \frac{3}{2} \approx 2^{\frac{7}{12}} = 1.498 \text{ (700c)}$$

The first is quite poor, and actually means the identification of C$^{\#}$ with C, while the second is significantly better, relating to the Pythagorean identification of E$^{\text{b}}$ and D$^{\#}$. In the equal tuning, these two approximations have the following meaning. The first approximation means that in the equal tempered scale with 7 tones, the just fifth is approximated by the 4^{th} of 7 tones. The second approximation means that in the equal tempered scale with 12 tones, the just fifth is approximated by the 7^{th} of 12 tones.

It is obvious that 1.498 approximates 1.5 better than 1.486 does. However, the question arises whether in the equal tempered scale with n tones the just fifth is even better approximated by the m-th tone for some $n \neq 12$. Mathematically, the problem boils down to calculating the fraction $\frac{m}{n}$ that best approximates the number x given by the condition:

$$\frac{3}{2} = 2^x \quad \text{i.e.} \quad x = \log_2 \frac{3}{2} = \frac{\log(3/2)}{\log 2}.$$

Since $\frac{\log(3/2)}{\log 2}$ is an irrational number, there is no fraction $\frac{m}{n}$ such that $x = \frac{m}{n}$. Irrational x can only be approximated by a fraction. The theory of continued fractions shows us how this can be done in the best way.

A continued fraction is a fraction of the form:

$$q_0 + \cfrac{1}{q_1 + \cfrac{1}{q_2 + \cfrac{1}{q_3 + \cfrac{1}{q_4 + \cfrac{1}{\dots}}}}}$$

$q_0, q_1, q_2, q_3, q_4, \dots$ are natural numbers, which we call the quotients of that continued fraction, and we write it more simply like this:

$$[q_0, q_1, q_2, q_3, q_4, \dots].$$

If the continued fraction has finitely many quotients q_i then it is truly a fraction, i.e. it is a rational number. For example,

$$[0,1,1,2] = 0 + \cfrac{1}{1 + \cfrac{1}{1+\frac{1}{2}}} = \frac{3}{5}$$

If the continued fraction has infinitely many quotients q_i then it is an irrational number. The reverse is also valid. Hence, every real number can be expressed in the form of a continued fraction, rational in finite form, and irrational in infinite form. For example,

$$\frac{7}{12} = 0 + \frac{1}{\frac{12}{7}} = 0 + \frac{1}{1+\frac{5}{7}} = 0 + \cfrac{1}{1+\cfrac{1}{1+\frac{2}{5}}} = 0 + \cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{1+\frac{1}{2}}}} = [0,1,1,1,2]$$

In the following example, it is important to note that:

$$\sqrt{2}-1=\frac{(\sqrt{2}-1)(\sqrt{2}+1)}{\sqrt{2}+1}=\frac{1}{\sqrt{2}+1}$$

Then we see that:

$$\sqrt{2}=1+\sqrt{2}-1=1+\frac{1}{\sqrt{2}+1}=1+\frac{1}{2+\sqrt{2}-1}=1+\frac{1}{2+\frac{1}{\sqrt{2}+1}}$$

$$=1+\frac{1}{2+\frac{1}{2+\sqrt{2}-1}}=1+\frac{1}{2+\frac{1}{2+\frac{1}{\sqrt{2}+1}}}=[0,1,2,2,\dots]$$

If we represent the real number x in the form of a (finite or infinite) continued fraction $[q_0, q_1, q_2, q_3, q_4, \dots]$ then its initial segments are the following fractions:

$$x_0=q_0=\frac{q_0}{1}=\frac{m_0}{n_0}\qquad x_1=q_0+\frac{1}{q_1}=\frac{q_0q_1+1}{q_1}=\frac{m_1}{n_1}$$

$$x_2=q_0+\frac{1}{q_1+\frac{1}{q_2}}=\frac{q_2\,(q_0q_1+1)+q_0}{q_2q_1+1}=\frac{m_2}{n_2}\dots$$

They are the best rational approximations of x, up to the size of their denominator (which we assume is fully reduced). This means that some fraction m/n better approximates x, than does the fraction $x_k = m_k/n_k$, only if $n > n_k$. For example, the initial segments of an infinite continued fraction $\sqrt{2} = [1, 2, 2, 2, 2 \dots]$, are the following fractions:

$$[1]=\frac{1}{1}\qquad[1,2]=\frac{3}{2}\qquad[1,2,2]=\frac{7}{5}\qquad[1,2,2,2]=\frac{17}{12}\text{ etc.}$$

We can express it more clearly as follows:

1	2	2	2	...
$\frac{1}{1}$	$\frac{3}{2}$	$\frac{7}{5}$	$\frac{17}{12}$	...

So, 7/5 is the best approximation of $\sqrt{2}$ with a denominator up to 5, and 17/12 is the best approximation with a denominator up to 12. It is possible that with a denominator between 5 and 12 there is an approximation better than 7/5 and worse than 17/12. It must be between $[1, 2, 2] = 7/5$ and $[1, 2, 2, 2] = 17/12$. From theory of continued fractions we know that the only candidate for such an approximation is $[1, 2, 2, 1] = 10/7$. Since 10/7 is a worse approximation than 7/5 (which can easily be checked on a computer) it follows that 17/12 is the first better approximation of $\sqrt{2}$ after 7/5.

More generally, if between the approximations $[q_0, q_1, q_2]$ and $[q_0, q_1, q_2, 4]$ there is one that is better than the first one, it must be of the form

$$[q_0,q_1,q_2,1],\quad[q_0,q_1,q_2,2]\quad\text{or}\quad[q_0,q_1,q_2,3].$$

We can apply these results about continued fractions to solving our problem of the optimal number of tones in an octave. We have shown that it boils down to finding the fractions that best approximate $x = \frac{\log(3/2)}{\log 2}$. This number should first be expressed in the form of a continued fraction:

$$\frac{\log\left(\frac{3}{2}\right)}{\log 2} = 0 + \frac{1}{\frac{\log 2}{\log\left(\frac{3}{2}\right)}} = 0 + \frac{1}{\frac{\log\left(\frac{3}{2}\right)\left(\frac{2}{3}\right)2}{\log\left(\frac{3}{2}\right)}} = 0 + \frac{1}{1 + \frac{\log\left(\frac{4}{3}\right)}{\log\left(\frac{3}{2}\right)}} = 0$$

$$+ \frac{1}{1 + \frac{1}{\frac{\log\left(\frac{4}{3}\right)\left(\frac{3}{4}\right)\left(\frac{3}{2}\right)}{\log\left(\frac{4}{3}\right)}}} = 0 + \frac{1}{1 + \frac{1}{\frac{\log\left(\frac{9}{8}\right)}{\log\left(\frac{4}{3}\right)}}} = 0 + \frac{1}{1 + \frac{1}{1 + \frac{1}{2 + \dots}}}$$

Continuing the process, we would get:

$$\frac{\log\left(\frac{3}{2}\right)}{\log 2} = [0, 1, 1, 2, 2, 3, 1, 5, \dots]$$

The corresponding approximations are as follows:

0	1	1	2	2	3	1	5	...
$\frac{0}{1}$	$\frac{1}{1}$	$\frac{1}{2}$	$\frac{3}{5}$	$\frac{7}{12}$	$\frac{24}{41}$	$\frac{31}{53}$	$\frac{179}{306}$	...
			$\frac{2}{3}$	$\frac{4}{7}$	$\frac{17}{29}$	...	...	...
					$\frac{10}{17}$	...	...	...

It can be checked on the computer that 10/17 is a worse approximation than 7/12, so it is has to be rejected. Hence, the best approximations of the number x, up to their denominator size, are:

$$\frac{3}{5} \quad \frac{4}{7} \quad \frac{7}{12} \quad \frac{17}{29} \quad \frac{24}{41} \quad \frac{31}{53} \quad \dots$$

The 17/29 approximation means that the octave should be divided into 29 microtones, with the 17^{th} tone being a fifth. The amount of that perfect fifth is $2^{17/29} = 1.5013$. The disadvantage of this division (as well as that of $41, 53, \dots$ micro-tones) is that 29 micro-tones in an octave, for any performance except electronic, is completely inappropriate. On the other hand, the fifth 4/7 (and even more 3/5 and 2/3) is quite bad because $2^{4/7} = 1.486 = 686$ cents, which is 17 cents away from a perfect fifth. As the optimum, we are left with 12 tones in an octave with the 7^{th} tone as a fifth. Namely, $2^{7/12} = 1.498 = 700$ cents, which is only 2 cents away from a perfect fifth.

We can conclude that equal temperament scale with correct fifths and an accessible number of tones can only have 12 semitones. This is a theoretically necessary choice, which musical practice reached even without theory.

We might ask ourselves what is the optimal equal temperament scale that has accurate fourths and an accessible number of semitones. The problem is mathematically reduced to finding the fractions that best approximate $\frac{\log(4/3)}{\log 2}$. Its continued fraction and its initial segments are:

$$\frac{\log\frac{4}{3}}{\log 2} = [0, 2, 2, 2, 3, \dots]$$

0	2	2	2	3	...
$\frac{0}{1}$	$\frac{1}{2}$	$\frac{2}{5}$	$\frac{5}{12}$	$\frac{17}{41}$	...
			$\frac{3}{7}$	$\frac{12}{29}$	...
				$\frac{7}{17}$	...

It can be verified on the computer that 7/17 is a worse approximation than 5/12, so it has to be rejected. The best approximations, up to their denominator size, are:

$$\frac{2}{5} \quad \frac{3}{7} \quad \frac{5}{12} \quad \frac{12}{29} \quad \frac{17}{41} \quad \ldots$$

We immediately see that 5/12 is the best approximation, with an accessible number of semitones. However, this is the same distribution that we found for correct fifths. In it, the fourth is the 5^{th} of 12 semitones, while the fifth is the 7^{th}.

If we want correct thirds, we will get somewhat different results. Namely, we have the following approximations for major and minor thirds:

$$\frac{\log \frac{5}{4}}{\log 2} = [0, 3, 9, \ldots] \qquad \frac{\log \frac{6}{5}}{\log 2} = [0, 3, 1, 4, \ldots]$$

0	3	9	...
$\frac{0}{1}$	$\frac{1}{3}$	$\frac{9}{28}$	...
	$\frac{1}{2}$	$\frac{8}{25}$	...
	$\frac{1}{1}$	$\frac{7}{22}$	...
		$\frac{6}{19}$	...
		$\frac{5}{16}$	...
		...	...

0	3	1	4	...
$\frac{0}{1}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{5}{19}$	...
		$\frac{1}{2}$	$\frac{4}{15}$	...
		$\frac{1}{1}$	$\frac{3}{11}$	...
			$\frac{2}{7}$	...

It can be verified on the computer that 6/19, 5/16 and another four fractions in 3^{rd} column on the left are worse approximations than 9/28, so they have to be rejected. Hence, the best octave division for correct major thirds is a division into 19 semitones, with the major third as the 6^{th} tone (one could think about 22, 25 or 28 semitones, but they will not be favourable for minor thirds).

It can be demonstrated on the computer that the approximation of 2/7 is less accurate than 5/19, so it has to be rejected. Hence, the best octave division for correct minor thirds is a division into 11, 15 or 19 semitones.

Considering that major thirds prefer 19 semitones, the 19-tone scale is definitely the best scale for thirds. In such a scale, an excellent minor third (316 cents) would be the 5^{th} semitone, a tolerable major third (379 cents) would be the 6^{th} semitone, a tolerable fifth (695 cents) would be the 11^{th} semitone, and a tolerable fourth

(505 cents) would be the 8^{th} semitone. Such a scale can be realized on ordinary keyboards by dividing the black keys into two and by inserting one new black key between E and F and H and c.

However, 19 semitones may still be too many, so the 12 magnificents rules almost all music (with the exception of contemporary microtonal music)

References

[1] Z. Šikić, *Matematika i muzika*, HMD, Zagreb 1999.
[2] Z. Šikić, *Generalized Pythagorean Comma*, in: CroArtScia2011 - Symmetry: Art&Science, 169–174, 2014.